

Exhibit A

Tissue-specific Expression of Unique mRNAs That Encode Proglucagon-derived Peptides or Exendin 4 in the Lizard*

(Received for publication, August 1, 1996, and in revised form, November 8, 1996)

Yuqing E. Chen† and Daniel J. Drucker‡

From the Department of Medicine, Banting and Best Diabetes Centre, The Toronto Hospital, University of Toronto, Toronto, Ontario, Canada M5G 2C4

Glucagon-like peptide 1 stimulates insulin secretion and inhibits glucagon secretion, gastric emptying, and feeding, suggesting it may be biologically useful for the treatment of diabetes. A lizard glucagon-like peptide 1 (GLP-1)-related peptide, exendin 4, binds to the GLP-1 receptor and mimics the actions of GLP-1 *in vivo*. To determine the genetic relationship between exendin 4 and GLP-1, we analyzed the structure and expression of pancreatic and intestinal proglucagon mRNAs in the reptile *Heloderma suspectum*. Two different proglucagon cDNAs (lizard proglucagon I (LPI) and lizard proglucagon II (LPII)), with unique 3'-untranslated regions were identified. Two LPI mRNA transcripts, ~1.8 and 2.1 kilobases, encoded glucagon and GLP-1 but not GLP-2 and were restricted in expression to the pancreas. In contrast, a 1.1-kilobase LPII mRNA transcript, encoding glucagon, GLP-1, and GLP-2 utilized a different 3'-untranslated region and was expressed in both pancreas and intestine. Lizard proglucagon mRNA transcripts were not detectable by reverse transcription-polymerase chain reaction or Northern blotting in salivary gland. A single class of lizard salivary gland proexendin cDNAs encoded the sequence of exendin 4 and a 45-amino acid exendin NH₂-terminal peptide. Exendin mRNA transcripts were expressed in the salivary gland, but not pancreas or intestine. These data demonstrate that GLP-1 and exendin 4 represent related yet distinct peptides encoded by different genes in the lizard.

The mammalian proglucagon gene is expressed in a restricted tissue-specific fashion in the A cells of the pancreatic islets, the L-cells of the intestinal mucosa, and in the central nervous system (1–3). In mammals, a single copy proglucagon gene gives rise to one mRNA transcript transcribed from an identical promoter in the pancreas, intestine, and brain (4–6). Posttranslational processing of proglucagon is highly tissue-specific, resulting in the release of different proglucagon-derived peptides (PGDPs)¹ from pancreas, intestine, and brain (1,

7, 8). In the pancreas, glucagon is the major bioactive PGDP, whereas in the intestine, processing of proglucagon to GLP-1, GLP-2, oxyntomodulin, and glicentin results in the liberation of a broader spectrum of bioactive peptides (8).

The PGDPs exhibit an increasing number of important biological activities. Glucagon, a 29-amino acid peptide hormone, is a key regulator of carbohydrate, protein, and lipid metabolism. Truncated forms of GLP-1, either the (7–37) or (7–36)-amide molecules, display potent glucose-dependent insulinotropic actions both *in vitro* and *in vivo* (9) and mice with a targeted disruption of the GLP-1 receptor gene develop diabetes (10). GLP-1 infusion in human subjects lowered postprandial glycemia and decreased insulin requirements in patients with both insulin-dependent diabetes mellitus and non-insulin-dependent diabetes mellitus, raising the possibility that GLP-1 may be a useful adjunct for the treatment of patients with diabetes (11). A role for GLP-1 in the central control of feeding behavior and satiety was recently suggested by experiments showing ICV administration of GLP-1 potentially inhibited feeding in rats (12). Taken together, these observations provide increasing evidence for the importance of GLP-1 in the physiological regulation of complex biological systems.

The isolation of a novel peptide from lizard venom that displays 52% identity to mammalian GLP-1 prompted studies examining the biological properties of this protein, designated exendin 4 (13). Experiments employing synthetic exendin 4 provided evidence that this peptide shares similar biological activities with mammalian GLP-1. Exendin 4 and truncated GLP-1(7–36)-amide increased cAMP levels in guinea pig acinar cell preparations (14), and exendin 4 was subsequently shown to bind the GLP-1 receptor, stimulate glucose-dependent insulin secretion, and increase both cAMP accumulation and insulin gene expression in cultured islet cell lines *in vitro* (15). The insulinotropic GLP-1-like properties of exendin 4 suggest that this lizard peptide may also be therapeutically efficacious for the treatment of patients with diabetes.

A smaller peptide derived from exendin 4 (following deletion of the first 8 amino acids), designated exendin-(9–39), is an effective antagonist of the GLP-1 receptor, blocking both the binding and biological activities of GLP-1 and exendin 4 (13, 15). Exendin-(9–39) has proven to be a physiologically relevant antagonist for assessing the biological importance of GLP-1 by blocking the binding and activity of GLP-1 *in vivo* (16, 17). The striking homology between lizard exendin 4 and mammalian GLP-1 raises the possibility that exendin 4 may simply represent the reptilian equivalent of GLP-1 that is also expressed in

* The work was initially supported in part by a grant from the Canadian Diabetes Association and subsequently by the Medical Research Council of Canada. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

The nucleotide sequence(s) reported in this paper has been submitted to the GenBank™/EBI Data Bank with accession number(s) U77611–U77613.

† Supported by a postdoctoral fellowship from the Juvenile Diabetes Foundation International.

‡ Supported by a Scientist Award from the Medical Research Council of Canada. To whom correspondence should be addressed: The Toronto Hospital, 206 Elizabeth St., CCRW3-838, Toronto, Ontario M5G 2C4. Tel.: 416-540-4125; Fax: 416-978-4108; E-mail: d.drucker@utoronto.ca.

¹ The abbreviations used are: PGDPs, proglucagon-derived peptides;

bp, base pair(s); PGDPs, proglucagon-derived peptides; cAMP, cyclic AMP; ENT, exendin NH₂-terminal peptide; GLP-1, glucagon-like peptide 1; GLP-II, glucagon-like peptide 2; kb, kilobase(s); LPI, lizard proglucagon I; LPII, lizard proglucagon II; ORF, open reading frame; RT-PCR, reverse transcription-polymerase chain reaction; UTR, untranslated region; IP-1, intervening peptide 1.

the lizard salivary gland. If this is correct, the exendin 4 sequence should be coencoded with glucagon and other PGDPs within the lizard proglucagon prohormone. Alternatively, exendin 4 could be derived from a novel gene and may represent a new member of a family of GLP-1-related peptides that exert biological activities comparable with GLP-1, in reptiles as well as in higher mammals. To elucidate the genetic relationship between exendin 4 and GLP-1, we isolated and characterized a series of novel *Heloderma suspectum* proglucagon and exendin cDNAs from the lizard.

EXPERIMENTAL PROCEDURES

Materials—Restriction enzymes, T4 DNA ligase, and DNA polymerase were from Pharmacia Biotech Inc. [α - 32 P]dATP (>800 Ci/mmol) was from ICN Radiochemicals (Toronto, Ontario). [α - 32 S]dATP (>600 Ci/mmol) was from Amersham Corp. All chemicals were from Sigma (Toronto, Ontario). *H. suspectum* lizard tissues were obtained from Dr. R. Murphy, Royal Ontario Museum (Toronto, Ontario). Two oligonucleotide primers (LP(A) and LP(B)) were a gift of Dr. D. M. Irwin, the Toronto Hospital (Toronto, Ontario). All other oligonucleotides for PCR and DNA sequencing, and for cloning lizard exendin, LE(A) and LE(B), were synthesized by ACGT Inc. (Toronto, Ontario).

RT-PCR—Total cellular RNA from lizard intestine, pancreas, and salivary gland was isolated by the acid guanidinium thiocyanate method (18). For first strand cDNA synthesis, 5 μ g of total cellular RNA was reverse-transcribed using the SuperScript preamplification system (Life Technologies, Inc.). An aliquot of the first strand cDNA (0.1 volume) was subsequently amplified by PCR (35 cycles, annealing temperature of 50 °C, extension temperature of 72 °C). The degenerate oligonucleotides were designed according to the conserved NH₂-terminal sequence (HSQQTFF) of mammalian, bird, and amphibian sequences: LP(A), 5'-GGCAATT CAY TCN CAR GGN ACN TT-3', where N is any nucleotide, R is purine, and Y is pyrimidine and the conserved NH₂-terminal sequence (HADGT) of GLP-1: LP(B), 5'-GAGGATCC NGT NCC RTC NGC RTG-3'. For isolation of lizard exendin, the degenerate primers were: LE(A), 5'-GCGGAATT CAY GGN GAR GGN ACN TT-3' and LE(B), 5'-CGCGGA TCC RTT YTT NAR CCA YTC-3'. Amplification products were cloned into the TA cloning vector pCRII from Invitrogen (San Diego, CA) and sequenced.

Isolation of Pancreatic Proglucagon cDNA Clones—A *H. suspectum* cDNA library, made from poly(A)⁺ mRNAs isolated from the pancreas of adult lizard, consisting of 1×10^6 independent recombinant phagemids, was constructed in the EcoRI sites of a pcDNA II vector (Invitrogen, San Diego, CA). The cDNA library was screened with a 209-bp EcoRI/BamHI [α - 32 P]dATP-labeled lizard cDNA generated by RT-PCR (see Fig. 1C). Nine positive clones were isolated and sequenced by using 32 S-dATP and a Sequenase version 2.0 kit (U. S. Biochemical Corp.) for direct dideoxynucleotide sequencing. In some cases, particular sequences were confirmed by subcloning DNA fragments into pBluescript II SK⁺ for dideoxynucleotide sequencing (19).

Isolation of Intestinal Proglucagon cDNA Clones—Two pairs of proglucagon gene primers were designed following isolation of the lizard pancreatic cDNAs. The 5' primer, designated LizGlu5 is: 5'-CTT CTC CAT ATC CCT TC-3' (nucleotides 47–63 in LPI; nucleotides 57–73 in lizard II, see Figs. 2B and 3B). The 3' primers are LizGlu3-1, 5'-ATA GTA GCA ACA GAT GG-3' (complementary to nucleotides 540 to 556 in LPI, Fig. 2B) and LizGlu3-2, 5'-ACA CAA CGT TAT GCT TC-3' (complementary to nucleotides 780–796 in LPII, Fig. 3B). RT-PCR was performed using lizard intestinal RNA as described above and these sets of primers. A ~700-bp band was amplified using LizGlu5 and LizGlu3-2 primers. The PCR product was subcloned and sequenced on both strands.

Isolation of Exendin cDNA Clones—A *H. suspectum* cDNA library, made from poly(A)⁺ mRNAs isolated from the salivary glands of an adult lizard, consisting of 1×10^6 independent recombinant phagemids, was constructed in the EcoRI sites of the pcDNA II vector (Invitrogen, San Diego, CA). The cDNA library was screened with a 95-bp EcoRI/BamHI [α - 32 P]dATP-labeled lizard cDNA generated by RT-PCR (see Fig. 6C).

Southern Blot Analysis—DNA was isolated from *H. suspectum* liver tissue and digested with various restriction enzymes. The digested DNA was electrophoresed in a 0.7% agarose gel, vacuum-blotted, and UV-fixed using a UV Stratalinker (Stratagene, San Diego, CA). The membrane was prehybridized in a solution containing 10% formamide, 3 \times SSPE (1 \times SSPE = 0.18 M NaCl, 0.01 M Na₂HPO₄, 1 mM EDTA), 1% SDS, 0.5% Biotite, and 0.5 mg/ml denatured herring sperm DNA for 4 h

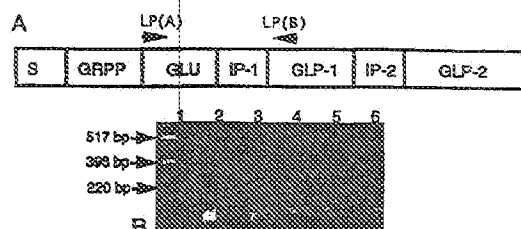


FIG. 1. Generation and analysis of *H. suspectum* proglucagon cDNAs. A, schematic representation of the location of two degenerate proglucagon primers (arrowheads) used for RT-PCR. S, signal peptide; GRPP, glucagon-related pancreatic polypeptide; GLU, glucagon; IP-1, intervening peptide 1; GLP-1, glucagon-like peptide 1; IP-2, intervening peptide 2; GLP-2, glucagon-like peptide 2. B, analysis of lizard proglucagon cDNAs generated by RT-PCR. Agarose gel electrophoresis of RT-PCR products generated using primers LP(A) and LP(B) from lizard pancreas (lane 2), intestine (lane 3), liver (lane 4), salivary glands (lane 5), and H₂O (lane 6) are shown along side DNA markers after electrophoresis in 2.5% agarose gel stained with ethidium bromide. The PCR product encodes 209 nucleotides of lizard proglucagon cDNA sequence.

at 55 °C. Hybridization was performed in a similar solution containing 10% dextran sulfate and probe A for proglucagon: a 201-bp [α - 32 P]dATP-labeled lizard proglucagon cDNA (nucleotides 243–443 in LPI and nucleotides 253–453 in LPII, see Figs. 2 and 3), or an exendin cDNA probe, for 12 h at 55 °C. The final wash was done at 60 °C. The membranes were exposed to Kodak XR-Omat RP-1 x-ray film with an intensifying screen at -70 °C. DNA fragment sizes were estimated using 1-kb DNA markers.

Northern Blot Analysis—RNA from different tissues was separated in 1.0% agarose gels containing 2.6 M formaldehyde and vacuum-blotted to a Zeta-Probe membrane. Blots were hybridized with the following proglucagon cDNA probes: A, a 201-bp lizard proglucagon cDNA as described above; B, a 861-bp 3'-UTR fragment from LPI; or C, a 340-bp 3'-UTR fragment from the LPII. The exendin blot in Fig. 6 was hybridized with a full-length lizard salivary gland exendin cDNA probe. Hybridization and washing conditions were carried out as described previously (3). The approximate sizes of various mRNA species were determined from the mobility of a coelectrophoresed 0.24–9.49-kb RNA ladder (Life Technologies, Inc.).

RESULTS

Isolation of a Lizard Proglucagon cDNA—As a first step in the elucidation of the potential relationship between exendin 4 and GLP-1, we generated a partial lizard proglucagon cDNA by RT-PCR. Two degenerate oligonucleotides (see "Experimental Procedures") were used to amplify first strand cDNAs prepared using RNA isolated from *H. suspectum* intestine, pancreas, and salivary gland (Fig. 1, A and B). Analysis of PCR products generated using cDNA from lizard pancreas revealed a ~215-bp partial proglucagon cDNA (Fig. 1B, lane 2). An identical albeit much less abundant band that comigrated with the pancreatic cDNA product was also obtained using the same primers and first strand cDNA from lizard intestine (Fig. 1B, lane 3). In contrast, no PCR products were generated using the same primers and first strand cDNA prepared from either lizard liver or salivary gland (Fig. 1B, lanes 4 and 5). The size of the PCR product was larger than one would predict (for a partial mammalian proglucagon cDNA using the identical primers) and raised the possibility that one of the lizard PGDPs, possibly IP-1, was larger than mammalian IP-1s, consistent with the results recently described for the chicken (20).

To ascertain the identity of the putative partial proglucagon cDNAs generated from pancreas and intestine, the PCR products were subcloned and subjected to DNA sequence analysis. Lizard pancreatic and intestinal (partial) cDNA clones contained an open reading frame encoding identical nucleotide sequences corresponding to: lizard glucagon, a larger than expected peptide immediately carboxyl-terminal to glucagon des-

TABLE I
Alignment of amino acid sequences of glucagon and glucagon-like peptides

Alignment of the amino acid sequences of lizard, trout, chicken, and human glucagon, GLP-1, and GLP-2 peptides. For comparative purposes, the sequence of lizard exendin is shown, with amino acid sequences similar to lizard GLP-1 in bold.

				% identity
Glucagon				
Lizard	HSQGTFTSDY	SKYLDTRRAQ	DPVQWLNT	
Trout I	--E--SN--	--QEE-M--	-----S	72.4
Chicken	-----	-----S	-----S	92.4
Human	-----	-----S	-----	98.5
Exendin and truncated glucagon-like peptide-1-(7-37)				
Exendin	HGEGTFTSDI	SKQNEEAVR	LFIEHLKRG	PSSGAPPPSG
Lizard	HADGRYTSOI	SSYLEGQAAK	EFIAWLNGR	G
Trout I	---T---V	-T--QD---	D-VS--KS--	A
Chicken	---E-T---	T-----	-----	-
Human	---E-TF---V	-----	-----K--	-
Glucagon-like peptide-2				
Lizard	HADGTFTSDY	NQLLDDIATQ	EFLKWLINQK	VTQ
Trout I	-V--S---V	-KV--SL-AK	-Y-L-VMTS-	TSG
Chicken	---S---I	-KI--M-AK	-----T-	---
Human	---S-SDEN	-TI--NL-AR	D-IN--QT-	I-D

quenced on both strands), was 1,050 bp and contained 95 bp of 5'-UTR, a 615-bp ORF, and 340 bp of 3'-UTR (Fig. 3B). The structure of the LPII cDNA corresponds to the mammalian proglucagon cDNAs (and the fish and chicken intestinal proglucagon cDNAs) that also contain IP-2 and GLP-2 (2, 5, 22, 23).

Comparison of the two classes of lizard pancreatic proglucagon cDNAs reveals that the nucleotide sequences from 1 to 531 in LPI are identical to the nucleotide sequences from positions 11 to 541 in the LPII cDNAs, resulting in the first 149 amino acids of the LPI and LPII proglucagons being identical, to end of the GLP-1 sequence (Figs. 2 and 3). Immediately 3' to the GLP-1 sequence, the sequences of LPI and LPII diverge, with LPI containing a single glutamic acid residue followed by a stop codon (Fig. 2B), whereas the open reading frame of LPII encodes a second intervening peptide (IP-2) and GLP-2 (Fig. 3B). Both LPI and LPII cDNAs contain a similar Kozak sequence at the beginning of the translation start sites (24). Two putative polyadenylation signals are found in the LPII cDNA (at nucleotides 719-724 and 825-830; Fig. 3B).

As neither LPI or LPII cDNAs contained an exendin 4-like sequence, we hypothesized that exendin 4 may be encoded within a unique lizard intestinal proglucagon cDNA that differed from LPI and LPII, possibly generated by alternative RNA splicing. A series of RT-PCR experiments were carried out using lizard intestinal RNA and two different combinations of lizard proglucagon primers; both primer sets utilized a single common 5'-primer (see "Experimental Procedures") and one of two primers corresponding to the unique 3'-UTRs (see "Experimental Procedures") from either LPI or LPII (Fig. 4). A single ~700-bp band was amplified using the LPII-specific primers. Sequence analysis of the cloned intestinal proglucagon cDNA products determined that the open reading frame encoded by the intestinal cDNA was identical to the sequence of pancreatic LPII (that also encodes IP-2 and GLP-2). In contrast, no intestinal PCR product was generated using LP-1-specific primers (data not shown). These results demonstrate that proglucagon gene expression in lizard intestine differs from pancreas and gives rise to a single class of proglucagon mRNA transcript that includes PGDPs carboxyl-terminal to GLP-1, consistent with previous studies in fish and chicken (20, 21).

Analysis of Lizard Proglucagon Gene Expression.—To ascertain whether mRNA transcripts related to proglucagon could be detected in salivary gland, the likely site of exendin 4 expression, RNA was isolated from *H. suspectum* intestine, liver, pancreas, and salivary gland and analyzed by Northern blot-

ting with different lizard proglucagon cDNA probes. Three mRNA transcripts of ~1.1, 1.6, and 2.1 kb were detected in RNA from lizard pancreas (Fig. 4A). In contrast, only a single 1.1-kb mRNA transcript (that comigrated with the pancreatic transcript) was detected in RNA from lizard intestine with a cDNA probe containing sequences common to both LPI and LPII cDNAs (Fig. 4A). Furthermore, using a cDNA probe specific to the LPI 3'-UTR sequences, only the 1.6- and 2.1-kb mRNA transcripts were detected in RNA from lizard pancreas; however, no hybridization was observed using the LPI 3'-UTR probe and RNA from lizard intestine (Fig. 4B). Consistent with the results predicted from cDNA cloning, the 1.1-kb mRNA transcript (corresponding to LPII) was observed in RNA from both pancreas and intestine using a cDNA probe specific for LPII (Fig. 4C). In contrast, no hybridizing transcripts were detected in RNA isolated from either lizard salivary gland or liver with any of the three proglucagon cDNA probes (Fig. 4, A-C). The lack of hybridization using liver and salivary gland RNA was not attributable to any noticeable differences in the quality of the various RNA preparations (Fig. 4D).

Isolation of Exendin Salivary Gland cDNAs.—To isolate exendin cDNA sequences from lizard salivary gland RNA, we prepared degenerate oligonucleotide primers against the lizard exendin 4 amino acid sequence for use in RT-PCR experiments (Fig. 5A). RNA from lizard salivary gland generated a PCR product of the predicted size (Fig. 5B), and sequencing of this partial cDNA demonstrated that it encoded the sequence of exendin 4 (Fig. 5C). To isolate a full-length exendin cDNA, we prepared and screened a lizard salivary gland cDNA library with the partial exendin cDNA probe. A single class of exendin cDNA clones was isolated from the library that encoded exendin and a 45-amino acid peptide NH₂-terminal to the exendin sequence, designated exendin NH₂-terminal peptide (ENTP). The ENTP sequence was followed by a dibasic pair of amino acids, lysine and arginine, characteristic of prohormone convertase cleavage sites, and the 40-amino acid exendin sequence (Fig. 5D). No peptide sequences exhibiting homology to ENTP were found following search of the latest GenBank™ release.

Northern blot analysis carried out using lizard RNA demonstrated the presence of a single exendin mRNA transcript in RNA from salivary gland (Fig. 6A). In contrast, no exendin mRNA transcripts were detected in RNA from lizard liver, pancreas, or intestine. The complexity of organization of the lizard exendin and proglucagon gene(s) was examined by Southern blot analysis

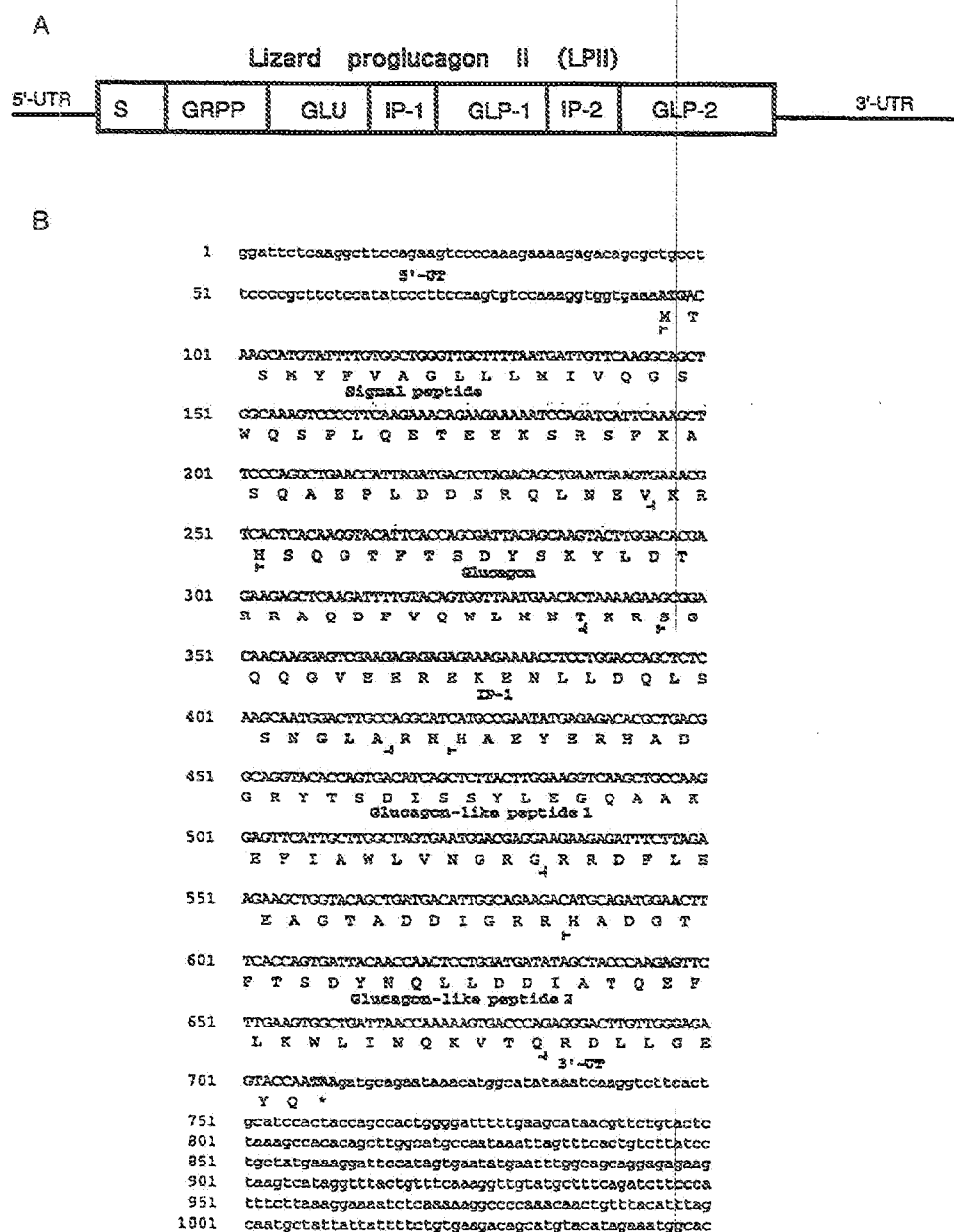


FIG. 3. A, structure of the LPII cDNA from *H. suspectum* pancreas. The abbreviations of PGDPs are the same as described in the legend to Fig. 1A. B, the nucleotide (open reading frame = *capital letters*; untranslated regions = *lowercase letters*) and derived amino acid sequences encoded by the LPII cDNA. The predicted ATG translation initiation and TAA stop codons are shown in *bold*. UT, untranslated region.

(Fig. 6, B and C). A simple pattern of hybridization was obtained using the exendin probe, clearly distinct from the bands generated with a lizard proglucagon cDNA probe common to the open reading frames of LPI and LPII. For example, a doublet of less than ~3.0 kb and a second fragment of ~1.8 kb was detected following *Eco*RI digestion (and hybridization) with the exendin probe. In contrast, hybridizing fragments of ~5.7 and 4.0 kb and a doublet of ~3.2 kb was detected with the proglucagon probe

(Fig. 6, B and C). These results are consistent with unique exendin and proglucagon genes and the presence of one, but most likely two *H. suspectum* proglucagon genes, as described previously for fish (21, 25, 26).

DISCUSSION

The biological diversity of the proglucagon-derived peptides in mammals derives from the cell-specific expression of prohormones.

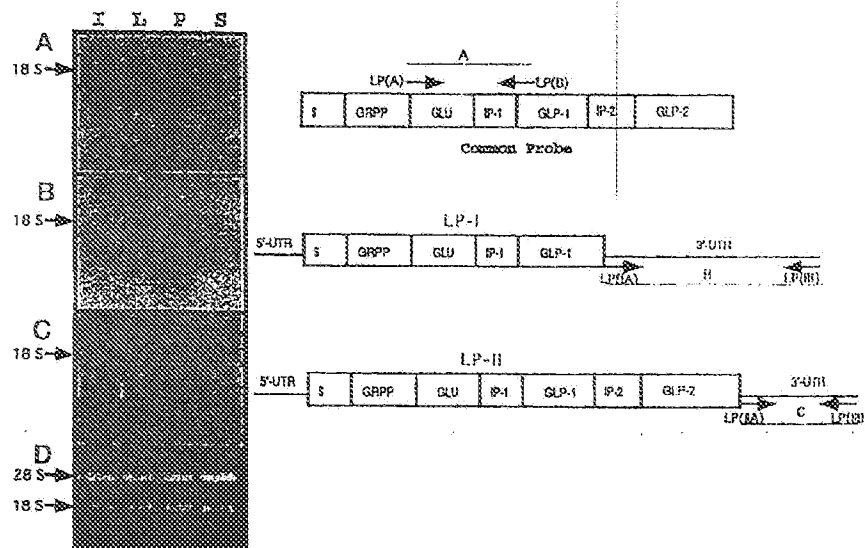


FIG. 4. Northern blot analyses of proglucagon mRNA transcripts in tissues from adult *H. suspectum* lizard. Total RNA, isolated from intestine (I), liver (L), pancreas (P), and salivary glands (S) was hybridized with the respective ³²P-labeled lizard proglucagon cDNAs, A-C, as shown to the right of each panel. The cDNA probes and location of primers used to generate probes used for each Northern analyses (labeled A-C) are shown. The agarose gel stained with ethidium bromide following electrophoresis is shown [a], and the migration position of 18 S ribosomal RNA is shown to the left of the panel (arrow).

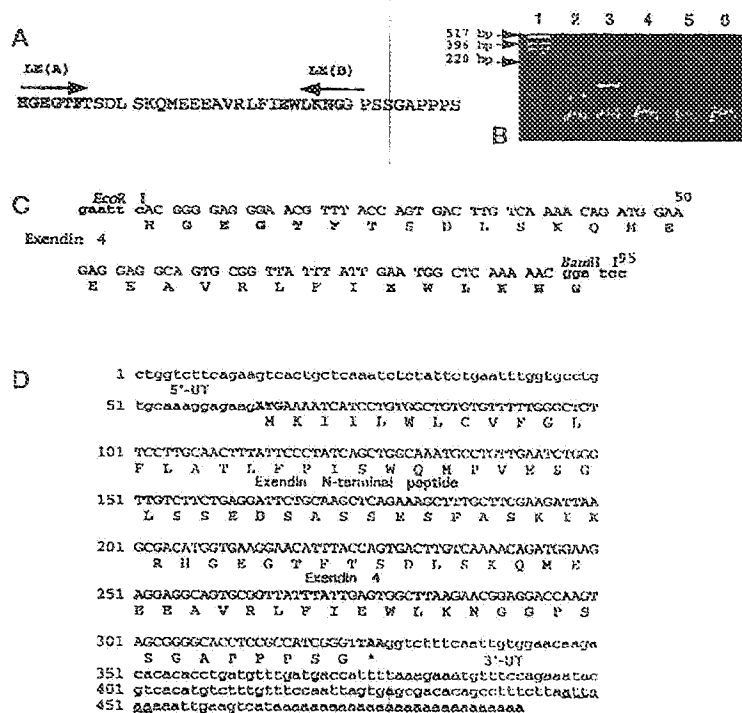


FIG. 5. A, strategy for RT-PCR isolation of exendin partial cDNAs from lizard salivary gland RNA. Amino acid sequences of exendin 4 and positions of primers LM(A) and LM(B) used for RT-PCR (arrows) are shown. B, results of RT-PCR using exendin 4 degenerate LEXAYL(B) primers showing positive PCR product in lane 3 (salivary gland RNA), but no exendin product in water alone, pancreas, intestine, or liver RNA (lanes 2, 4, 5, and 6, respectively). C, nucleotide and predicted amino acid sequence of cloned exendin RT-PCR cDNA product from lizard salivary gland. D, schematic representation of full-length proexendin cDNA isolated from the lizard salivary gland cDNA library. The translation start site and stop site are indicated in bold. UT, untranslated region. A putative polyadenylation consensus sequence is underlined.

more convertases that generate a unique profile of PGDPs in pancreas and intestine (27-30). Whereas a single proglucagon mRNA transcript (identical in pancreatic and intestinal tissues) has been detected to date in all mammals studied,

chicken and fish use alternative RNA splicing to generate a tissue-specific pattern of proglucagon mRNA transcripts that express different open reading frames in pancreas and intestine (20, 21). Furthermore, only a single class of proglucagon

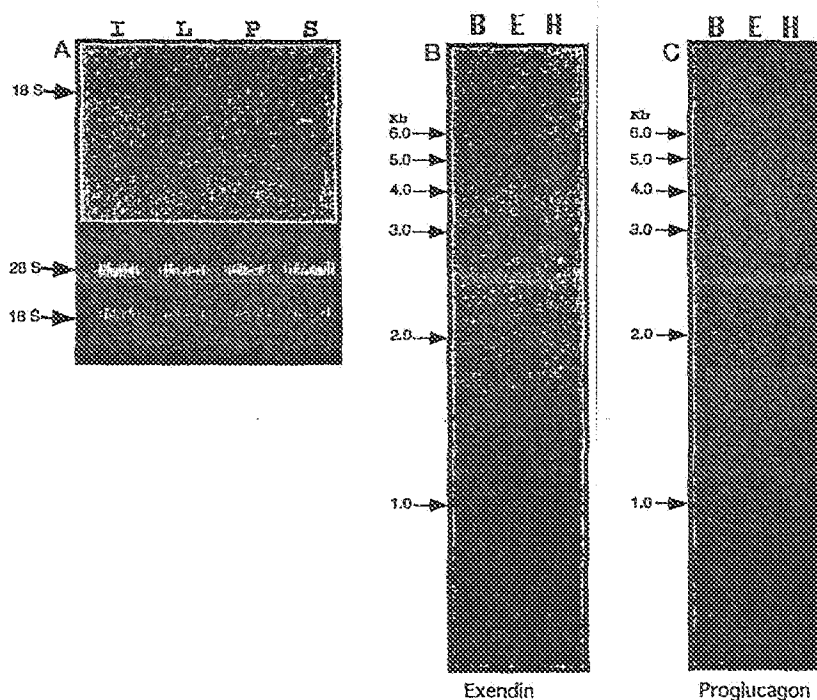


FIG. 6. A, Northern blot analysis of 20 µg of total cellular lizard RNA from intestine (I), liver (L), pancreas (P), and salivary gland (S) hybridized with the lizard exendin cDNA probe. The ethidium bromide-stained gel and the positions of 18 and 28 S ribosomal RNA are shown. B and C, Southern blot analysis of *H. suspectum* genomic DNA hybridized with a lizard exendin cDNA probe (B) or a lizard proglucagon cDNA probe common to both lizard I and lizard II cDNAs (C). *H. suspectum* genomic DNA (20 µg) was digested with *Bam*HI (B), *Kco*II (C), or *Hind*III (H). The relative sizes of the DNA fragments were determined from a coelectrophoresed DNA ladder shown at the left in kilobases.

mRNA transcript (and ORF) is present in pancreas and intestine from chicken or fish.

In contrast to proglucagon gene expression in mammals, chicken, or fish, the characterization of proglucagon gene expression in *H. suspectum* described here is clearly unique for several reasons. First, two structurally different classes of proglucagon mRNA transcripts, encoding different proglucagons and unique 3'-UTRs were isolated from lizard pancreas, but only a single type of proglucagon cDNA was obtained from lizard intestine. Furthermore, the 1.1-kb (LPII) transcript (that contains the largest ORF with proglucagon sequences COOH-terminal to GLP-1) is expressed in both pancreas and intestine, likely as a result of alternative RNA splicing. In contrast, the pancreas-specific unique 1.6- and 2.1-kb transcripts represent LPI cDNAs that encode glucagon and GLP-1, but not GLP-2 (Fig. 4). The likely explanation for the generation of three different pancreatic proglucagon mRNAs is alternative splicing of a single (or possibly two highly related) lizard proglucagon RNA(s) to generate mRNAs with different open reading frames and unique 3'-untranslated regions.

Although we did not isolate proglucagon cDNAs approximating 1.6 or 2.1 kb in size after screening the pancreatic cDNA library, the size of these two RNA transcripts may be attributable in part to either a long poly(A) tail, differential utilization of polyadenylation sites in the 3'-UTR, alternative RNA splicing (perhaps to include additional 3'-UTR sequence) of a mature transcript and/or incomplete processing of a nuclear RNA precursor (for the 2.1-kb RNA). Precedents for several of these possibilities derives from previous studies of proglucagon gene

expression in various species. A second smaller pancreatic proglucagon mRNA transcript was identified and characterized in a hamster islet cell line (31) that arises as a consequence of the utilization of a more proximal poly(A) consensus sequence in the hamster proglucagon 3'-UTR. Furthermore, different classes of pancreatic proglucagon mRNA transcripts (with identical open reading frames but longer poly(A) tails compared with wild-type pancreatic proglucagon RNAs) have been detected in RNA isolated from human and rat pancreatic endocrine tumors and fetal pancreas (32, 33), and alternative RNA processing likely accounts for tissue-specific expression of proglucagon mRNAs in fish and chicken (20, 21). Although the precise structural basis for the difference in size of the 1.6- and 2.1-kb lizard pancreatic proglucagon mRNAs remains to be determined, the results of our RT-PCR, Northern blotting and cDNA cloning experiments strongly suggest that these two LPI mRNAs likely share common open reading frames and at least some 3'-UTR sequence, in contrast to the different ORF and 3'-UTR of the LPII mRNA. Furthermore, the nucleotide sequences of the LPI and LPII cDNAs extending up to GLP-1 are identical. Hence LPI and LPII may be derived from the same gene and alternative splicing accounts for the carboxyl-terminal differences, or alternatively, they may represent the products of two highly related genes.

An initial goal of these studies was the clarification of the possible relationship between exendin 4 and lizard GLP-1. Although exendin 4, as isolated and sequenced from *H. suspectum* venom, is 52% identical to mammalian GLP-1 (13), the cloning of lizard proglucagon cDNAs demonstrated that the

sequence of exendin 4 is not encoded by lizard preproglucagon, and lizard exendin 4 is now shown to be only 45% identical to lizard GLP-1. The isolation of exendin 4 cDNA clones from lizard salivary gland, taken together with the results of Northern and Southern blot analyses, clearly establish for the first time that exendin 4 and lizard GLP-1 are unique peptides encoded by different genes.

Lizard GLP-1 shares 90 and 84% identity to chicken and human GLP-1, respectively, and lizard glucagon shares 92 and 97% identity to chicken and human glucagon, respectively (Table 1). In contrast, although lizard and chicken GLP-2 are reasonably similar in sequence (76% identity), lizard GLP-2 is much less similar than the rat or human (46%) peptides. Furthermore, lizard and chicken IP-1s are both much larger than all other mammalian and fish IP-1 sequences. This also implies that lizard and chicken oxyntomodulin and glicentin, that both contain the IP-1 sequence at their carboxyl terminus, are considerably larger than their mammalian counterparts. Taken together, the differential conservation of peptide homologies raises the possibility that some of the lizard PGDPs may have diverged not only in sequence, but also in biological function, from their mammalian counterparts. Interestingly, although rat and human GLP-1 are separated from IP-1 by a pair of basic amino acid residues (KR), the cognate lizard sequence is RH. The absence of the dibasic amino acid pair at the NH₂ terminus of lizard GLP-1 raises the possibility that the amino terminus of lizard GLP-1 may be recognized and processed in a manner that differs from the processing observed for vertebrate GLP-1s. In contrast, the first 2 amino acids of truncated GLP-1, HA, are identical in the human and lizard. The functional implications of the NH₂-terminal differences in lizard GLP-1 sequence remain to be established.

Taken together, the demonstration that exendin 4 is encoded by a lizard gene distinct from proglucagon in the lizard raises the possibility that other species, perhaps mammals, may also contain distinct exendin genes. The observation that lizard exendin 4 binds to the mammalian GLP-1 receptor and stimulates insulin secretion, whereas a truncated exendin 4 (9–39) peptide functions as an antagonist of the mammalian GLP-1 receptor both *in vitro* and *in vivo* (34) provides indirect evidence supporting the existence of a mammalian exendin 4 gene with potential physiological implications for control of insulin secretion *in vivo*.

Acknowledgments—We express our gratitude to Dr. David M. Irwin for his helpful suggestions, Dr. R. Murphy for his help in procuring lizard tissues, and to Drs. M. Zhang and N. Wins for their technical assistance. We thank the members in Dr. Drucker's laboratory for their discussion.

REFERENCES

1. Mojsov, S., Heinrich, G., Wilson, I. B., Ravazzola, M., Orci, L., and Habener, J. F. (1988) *J. Biol. Chem.* 263, 11880–11889
2. Drucker, D. J., and Aja, S. (1988) *J. Biol. Chem.* 263, 13475–13479
3. Drucker, D. J., and Brubaker, P. L. (1989) *Proc. Natl. Acad. Sci. U.S.A.* 86, 3953–3957
4. Novak, U., Wilks, A., Busell, G., and McEwen, S. (1987) *Eur. J. Biochem.* 164, 553–558
5. Heinrich, G., Gros, P., Lund, P. K., Bentley, R. C., and Habener, J. F. (1984) *Endocrinology* 115, 2178–2181
6. Lee, Y. C., Brubaker, P. L., and Drucker, D. J. (1990) *Endocrinology* 127, 2217–2222
7. Orskov, C., Holst, J. J., Kauttsen, S., Baldissere, F. G. A., Poulsen, S. S., and Nielsen, O. V. (1988) *Endocrinology* 119, 1457–1475
8. Orskov, C., Holst, J. J., Poulsen, S. S., and Kirkegaard, P. (1987) *Diabetologia* 30, 874–881
9. Fehm, H.-C., Goke, R., and Goke, B. (1995) *Endocr. Rev.* 16, 390–410
10. Scroochi, L. A., Brown, T. J., MasLusky, N., Brubaker, P. L., Averbach, A. B., Joyner, A. L., and Drucker, D. J. (1998) *Nat. Med.* 2, 1254–1258
11. Gutniak, M., Orskov, C., Holst, J. J., Ahlert, B., and Efendic, S. (1992) *N. Engl. J. Med.* 326, 1316–1322
12. Turton, M. D., O'Shea, D., Gunn, I., Beak, S. A., Edwards, C. M. B., Meeran, K., Choi, S. J., Taylor, G. M., Heath, M. M., Lambert, P. D., Wilding, J. P. H., Smith, D. M., Ghatei, M. A., Herbert, J., and Bloom, S. R. (1996) *Nature* 379, 69–72
13. Eng, J., Kleiman, W. A., Singh, L., Singh, G., and Raufman, J.-P. (1992) *J. Biol. Chem.* 267, 7403–7406
14. Raufman, J.-P., Singh, L., Singh, G., and Eng, J. (1992) *J. Biol. Chem.* 267, 21432–21437
15. Goke, R., Fehm, H.-C., Linn, T., Schmidt, H., Krause, M., Eng, J., and Goke, B. (1993) *J. Biol. Chem.* 268, 19650–19655
16. Wang, Z., Wang, H. M., Owji, A. A., Smith, D. M., Ghatei, M. A., and Bloom, S. R. (1995) *J. Clin. Invest.* 96, 417–421
17. Dalesio, D. A., Vogel, R., Prigione, R., Lashchenko, E., Koerber, D., Eng, J., and Zinsack, J. W. (1995) *J. Clin. Invest.* 97, 133–138
18. Chomczynski, P., and Sacchi, N. (1987) *Anal. Chem.* 59, 156–159
19. Chen, E. Y., and Seeburg, P. H. (1985) *DNA (N.Y.)* 4, 165–170
20. Irwin, D. M., and Wong, J. (1996) *Mol. Endocrinol.* 9, 267–277
21. Lund, P. K., Goodman, R. H., Deo, P. C., and Habener, J. F. (1982) *Proc. Natl. Acad. Sci. U.S.A.* 79, 345–349
22. Bell, G. I., Santerre, R. F., and Mullenbach, G. T. (1983) *Nature* 305, 716–718
23. Lopez, L. C., Fraser, M. L., Su, C. J., Kumar, A., and Saunders, G. F. (1983) *Proc. Natl. Acad. Sci. U.S.A.* 80, 5485–5489
24. Kozak, M. (1986) *Cell* 44, 283–292
25. Lund, P. K., Goodman, R. H., and Habener, J. F. (1981) *J. Biol. Chem.* 256, 6515–6518
26. Lund, P. K., Goodman, R. H., Montminy, M. R., Deo, P. C., and Habener, J. F. (1983) *J. Biol. Chem.* 258, 3280–3284
27. Rouille, Y., Westermarck, G., Martin, S. K., and Steiner, D. F. (1994) *Proc. Natl. Acad. Sci. U.S.A.* 91, 3242–3246
28. Rouille, Y., Martin, S., and Steiner, D. F. (1995) *J. Biol. Chem.* 270, 26486–26496
29. Rothenberg, M. E., Ellertson, C. D., Klein, K., Zhou, Y., Lindberg, I., McDonald, J. K., Mackin, R. B., and Nee, B. D. (1995) *J. Biol. Chem.* 270, 10136–10145
30. Dhanvantari, S., Sekish, N. G., and Brubaker, P. L. (1996) *Mol. Endocrinol.* 10, 342–355
31. Lee, Y. C., and Drucker, D. J. (1990) *Mol. Endocrinol.* 4, 800–806
32. Philippe, J., Powers, A. C., Mojsov, S., Drucker, D. J., Comi, R., and Habener, J. F. (1988) *Diabetes* 37, 1647–1651
33. Philippe, J., Chick, W. L., and Habener, J. F. (1987) *J. Clin. Invest.* 79, 351–358
34. Raufman, J.-P. (1996) *Regul. Pept.* 61, 1–18